

# Exploiting Spatial Correlation in Energy Constrained Distributed Detection

Juan Augusto Maya, *Student Member, IEEE*, Cecilia G. Galarza and Leonardo Rey Vega, *Member, IEEE*.

**Abstract**—We consider the detection of a correlated random process immersed in noise in a wireless sensor network. Each node has an individual energy constraint and the communication with the processing central units are affected by the path loss propagation effect. Guided by energy efficiency concerns, we consider the partition of the whole network into clusters, each one with a coordination node or *cluster head*. Thus, the nodes transmit their measurements to the corresponding cluster heads, which after some processing, communicate a summary of the received information to the fusion center, which takes the final decision about the state of the nature. As the network has a fixed size, communication within smaller clusters will be less affected by the path loss effect, reducing energy consumption in the information exchange process between nodes and cluster heads. However, this limits the capability of the network of beneficially exploiting the spatial correlation of the process, specially when the spatial correlation coherence of the process is of the same scale as the clusters size. Therefore, a trade-off is established between the energy efficiency and the beneficial use of spatial correlation. The study of this trade-off is the main goal of this paper. We derive tight approximations of the false alarm and miss-detection error probabilities under the Neyman-Pearson framework for the above scenario. We also consider the application of these results to a particular network and correlation model obtaining closed form expressions. Finally, we validate the results for more general network and correlation models through numerical simulations.

**Index Terms**—Distributed detection; correlated measurements; energy and bandwidth constraints; wireless sensor networks; hierarchical clustering; spectrum sensing; cognitive radio.

## I. INTRODUCTION

**D**ISTRIBUTED detection based on wireless sensor networks (WSN) is a topic which has attracted great interest in recent years (see [1] and references therein). A typical WSN has a large number of sensor nodes which are generally low-cost battery-powered devices with limited sensing, computing, and communication capabilities. Sensors acquire noisy measurements, perform simple data processing and propagate the information into the WSN to reach a decision about a physical phenomenon occurring in the coverage area.

Network resources, such as energy and bandwidth, are scarce, expensive, and key variables when the design is focused on processing latency and detection performance. Clever data processing strategies are required to maximize

performance under resource constraints. Appropriate choices of the processing strategy could largely impact on the total cost or on the life cycle of the WSN, and as such, make its deployment on remote locations economically viable or not.

### A. Related work

Distributed detection theory has been much studied in the past. Starting with the seminal work of Tenney and Sandell [2], several results have been derived on how each node process the available information and communicates with the *fusion center* (FC) where the final decision on the true state of nature is taken. Under this setup, digital transmission schemes, where appropriate quantization rules have to be designed have been considered in [3]–[5] (see also references therein). On the other hand, analog communication schemes were also studied in the past (see for example [6]–[11]).

Instead of independent observations, often assumed in the previously mentioned works, we explore a more realistic scenario with spatial dependence measurements. Distributed detection problems with correlated observations can be considerably more challenging than their conditionally independent counterparts [12]. Moreover, new theoretical frameworks have been proposed recently [13], [14].

The design of distributed processing strategies exploiting spatial and/or temporal correlation among data is, in general, an open problem. It is well known that signal correlation can help to improve the detection performance, specially when low quality sensor measurements are available [8], [15], [16]. Clearly, a clever use of the correlation in space and/or time requires cooperation among nodes. This problem has also been studied in previous works [16]–[19]. Additional studies have explored the effects of correlation on the performance of distributed detection systems [20], [21].

An important application that has gained a lot of attention in the recent years under this scenario is spectrum sensing for cognitive radio [22], [23]. Here, unlicensed, or secondary users (SU), want to detect the presence or absence of the licensed, or primary users (PU), in order to use the spectrum when it is available. SUs could cooperate and build a WSN to achieve this task. When a PU uses the spectrum, it transmits a signal that may be sensed by the WSN. Due to the shadowing propagation effect, different close nodes sense correlated measurements. Thus, dependent observations must be considered in order to develop reliable spectrum sensing schemes. Moreover, the signal model and the scheme developed in this work can be directly applied to a cognitive radio scenario [16].

The authors are with the University of Buenos Aires (UBA), Buenos Aires, Argentina. L. Rey Vega and C. Galarza are also with the CSC-CONICET, Buenos Aires, Argentina. Emails: {jmay, lrey, cgalar}@fi.uba.ar

This work was partially supported by the Peruh grant (UBA), and the projects UBACYT 20020130100751BA and CONICET PIP 112 20110100997.

Typically, the nodes of a WSN are distributed on a vast area to sense the whole region of interest. Thus, the energy needed for reliable communication increases with the transmission distance due to the effect of path loss. In order to save transmission energy, it is convenient to reduce the region where sensors transmit high precision data. In this way, a hierarchical WSN is introduced: sensors are grouped in clusters with a designated cluster head (CH). Each CH processes the data independently of the other ones and communicates a summary of the information gathered by its nodes to the FC, where the final decision about the state of the nature is made. Previous works [6], [24] have already considered the idea of partitioning the region to be sensed into clusters. However, in [24] the authors considered conditionally independent and identically distributed measurements, and in [6], spatially coherence regions were assumed, but dependence among them is not taken into account.

### B. Contributions

In this paper, we develop a clustering WSN model with energy and bandwidth constraints. We assume that each cluster processes the information independently of other clusters. Moreover, we consider that each individual cluster only knows the local correlation structure of the process. This hypothesis simplifies the scheme in the sense that communication of high data rate information between CHs or nodes from different clusters is avoided. As these exchanges would require to overcome path loss effects, an energy saving in each node will be obtained. However, in a WSN with correlated measurements, this strategy limits the ability of the network of beneficially exploiting the correlation of the process. Therefore, a trade-off is established between the beneficial use of the underlying random process spatial correlation and the energy spent by each node in the network. This is a trade-off between the energy required to reliably communicate each node measurement and the ability to effectively use the cooperation scheme required for exploiting the spatial correlation information.

To characterize this trade-off, we derive tight approximations of the false alarm and miss-detection error probabilities for statistics built with sum of independent although not necessarily identically distributed random variables. The expressions involve the sum of the logarithmic moment generating function (LMGF) of each random variable and its first and second derivatives. For example, in this paper we consider Gaussian processes for which the statistic is a quadratic form of a Gaussian correlated random vector. In this case, the statistic can be expressed as a weighted sum of i.i.d. random variables and the error probabilities can be easily computed.

We obtain closed forms expressions for a particular model of the sensed process. Also, we numerically compute the average performance for more general processes and when the sensors are randomly placed in space and show that similar conclusions can be drawn. These conclusions are summarized as follows:

- 1) In a WSN with Gaussian measurements there are two clearly different operating regimes: an energy limited

regime, where the energy dedicated to transmit the data limits the performance, and a correlation-limited regime, where the performance is limited by the poor use of the spatial correlation of the process. The ranges of validity of each of these regimes are given by different parameters of the WSN some of which are under control of the designer. The characterization of this problem gives crucial information for the deployment of a WSN.

- 2) In some cases of interest the use of local correlation information is almost as good as the use of the global correlation statistics of the sensed processes. This is a very important issue from an implementation point of view, as the use of global correlation information requires a greater degree of complexity in the network.
- 3) The use of local correlation information becomes more profitable when the sensor measurements are less reliable (i.e., the measurement signal to noise ratio is low).

### C. Notation and Organization

Scalars are written in italic, vectors in boldface and matrices in capital letters.  $A_n$  and  $B_{nm}$  are square and rectangular matrices of sizes  $n \times n$  and  $n \times m$ , respectively. We define column vectors with components  $a_k$  as  $\mathbf{a} = [a_k]_{k \in \mathcal{A}}$ , where  $k$  belongs to the set of indexes  $\mathcal{A}$ . Similarly, we denote matrices  $A_{nm} = [\mathbf{a}_k]_{k \in \mathcal{A}}$  built through the horizontal concatenation of the column vectors  $\{\mathbf{a}_k\}$ , where  $\mathbf{a}_k \in \mathbb{C}^{n \times 1}$  and the cardinality of  $\mathcal{A}$  is  $|\mathcal{A}| = m$ .  $\det(A_n)$  and  $\text{tr}(A_n)$  are the determinant and the trace of  $A_n$ , respectively.  $(\cdot)^T$  and  $(\cdot)^H$  denote transpose and transpose conjugate. We do not distinguish between random variables and their realization values when the context is clear.  $p(\mathbf{x}|\mathcal{H})$  is the probability density function (p.d.f.) of  $\mathbf{x}$  under  $\mathcal{H}$ .  $\mathbb{P}_i(\cdot)$ , and  $\mathbb{E}_i(\cdot)$  denote probability and the expectation respectively, both computed under hypothesis  $\mathcal{H}_i$ . Given two statistics  $T$  and  $T'$ , we denote that they have the same error probabilities writing  $T \equiv T'$ . Given a function  $\mu(\cdot)$ , its first and second derivatives are denoted as  $\dot{\mu}(\cdot)$  and  $\ddot{\mu}(\cdot)$ . The notation  $f(n) = \mathcal{O}(g(n))$  means that  $\lim_{n \rightarrow \infty} f(n)/g(n) = c$ , where  $c$  is a non-zero constant.

The paper is organized as follows. In Section II, we present the WSN model and the binary hypothesis testing problem. In Section III, we describe the precoding strategies used for stationary and non-stationary random processes. In Section IV, we derive a new global statistic and formulate the main problem of the paper. In Section V, we compute tight approximations of the error probabilities to be used in Section VI to analyse the performance of the proposed strategy. Finally, in Section VII we elaborate on the main conclusions. Technical proofs are provided in the appendices.

## II. WSN MODEL

We analyse a sensor network with three hierarchies: nodes, clusters with their corresponding CH, and the FC. The nodes are grouped in clusters. Each CH gathers information from the nodes in its corresponding cluster, performs some processing independently of other clusters and transmit a summary to the

FC, where the final decision is made. A sketch of the WSN is shown in Fig. 1.

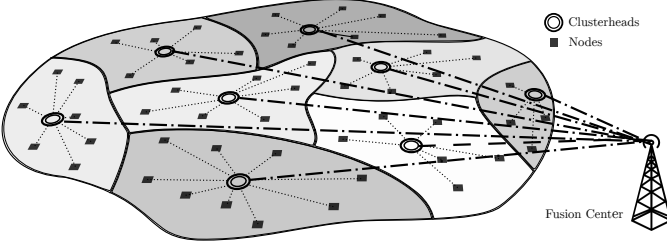


Fig. 1. A hierarchical WSN, where the information is propagated from the nodes to the CHs, and from them to the FC, which decides about the state of the nature.

The WSN has  $m$  clusters and a total amount of  $n$  nodes which are arbitrary distributed in an area  $nA_0$ , which for simplicity it will be assumed to be a square and  $A_0$  a constant value (with  $m^2$  units). Each cluster has  $l_i$  nodes,  $i = 1, \dots, m$ , such that  $n = \sum_{i=1}^m l_i$ . Let  $\mathcal{C}_i$  be the set of sensors of the  $i$ -th cluster,  $i = 1, \dots, m$ .  $\{\mathcal{C}_i\}_{i=1}^m$  is a partition of the set of sensors  $[1 : n]$ , i.e.,  $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset \forall i \neq j$ ,  $i = 1, \dots, m$ , and  $\bigcup_{i=1}^m \mathcal{C}_i = [1 : n]$ . In the following, we provide a description on the main elements of the general WSN model we will work with.

The  $k$ -th sensor of the WSN takes a measurement described under each hypothesis by

$$\begin{cases} \mathcal{H}_1 : & y_k = s_k + v_k, \\ \mathcal{H}_0 : & y_k = v_k, \end{cases} \quad k = 1, \dots, n, \quad (1)$$

where  $s_k$  is a correlated zero-mean circularly-symmetric complex Gaussian process with variance  $\sigma_s^2$  and autocorrelation function (ACF)  $E[s_j s_k] = R_s(j, k)$  with  $j, k \in [1 : n]$ . In general,  $\{s_k\}$  is a non-stationary spatial process. We also assume that  $v_k$  is a zero-mean circularly-symmetric complex white Gaussian noise independent of  $s_k$  with variance  $\sigma_v^2$ . Thus,  $y_k$  is Gaussian distributed either under  $\mathcal{H}_0$  or  $\mathcal{H}_1$ . The above model can be easily generalized to the case in which each sensor takes several measurements at different time instants, and the processes have temporal correlation. All the mathematical developments and conclusions in this paper can be extended to that case. However, in order to keep things simple we consider only the case in which each sensor takes only one measurement. This is in line with several works on distributed detection on WSN [8]–[10].

Define the vector  $\mathbf{y}_i = [y_k]_{k \in \mathcal{C}_i}$  as the set of measurements taken by the nodes in the  $i$ -th cluster. Define  $\mathbf{s}_i = [s_k]_{k \in \mathcal{C}_i}$  and  $\mathbf{v}_i = [v_k]_{k \in \mathcal{C}_i}$  accordingly. The covariance matrices of  $\mathbf{s}_i$  and  $\mathbf{v}_i$  are, respectively,  $\Sigma_{s, l_i}$  and  $\sigma_v^2 I_{l_i}$  where the elements of  $\Sigma_{s, l_i}$  are  $R_s(j, k)$ , with  $j, k \in \mathcal{C}_i$ , and  $I_{l_i}$  is the identity matrix of dimension  $l_i$ . Therefore, the covariance matrix of  $\mathbf{y}_i$  under  $\mathcal{H}_0$  is  $\Sigma_{0, l_i} = \sigma_v^2 I_{l_i}$ , and under  $\mathcal{H}_1$  is  $\Sigma_{1, l_i} = \Sigma_{s, l_i} + \sigma_v^2 I_{l_i}$ . Let  $\mathbf{y} = [\mathbf{y}_i]_{i \in [1:m]}$  be the measurements taken by all nodes in the network and define  $\mathbf{s} = [\mathbf{s}_i]_{i \in [1:m]}$ ,  $\mathbf{v} = [\mathbf{v}_i]_{i \in [1:m]}$ . The covariance matrices of  $\mathbf{s}$  and  $\mathbf{v}$  are, respectively,  $\Sigma_{s, n}$  and  $\sigma_v^2 I_n$  where the elements of  $\Sigma_{s, n}$  are  $(\Sigma_{s, n})_{j, k} = R_s(j, k)$ . The covariance matrix of  $\mathbf{y}$  under  $\mathcal{H}_0$

and  $\mathcal{H}_1$ , are, respectively,

$$\Sigma_{0, n} = \sigma_v^2 I_n, \quad \Sigma_{1, n} = \Sigma_{s, n} + \sigma_v^2 I_n. \quad (2)$$

Notice that  $\Sigma_{0, n}$  and  $\Sigma_{1, n}$  can be thought as  $m \times m$  block matrices, where the diagonal blocks are  $\Sigma_{0, l_i}$  and  $\Sigma_{1, l_i}$ ,  $i = 1, \dots, m$ .

### III. CLUSTER PRECODING STRATEGIES

In this work, we consider that the communication between the nodes and their corresponding CH is through either a *parallel-access channel* (PAC) or multiple channel uses<sup>1</sup> of a *multiple-access channel* (MAC). In the case of the PAC, the amount of channel uses coincides with the number of available orthogonal channels in each cluster. In both cases, we assume that the nodes communicate with its corresponding CH without any interference from other clusters. Consider now that the  $i$ -th cluster has  $l'_i$  channel uses available for transmission, with  $l'_i \leq l_i$ . Each channel use is associated to either an orthogonal time slot or a frequency band. Then, the processing strategy may use  $l'_i$  time slots,  $l'_i$  frequency bands or a combination of them such that the product of time slots and frequency bands is  $l'_i$ . An illustration of the distributed scheme is shown in Fig. 2(a).

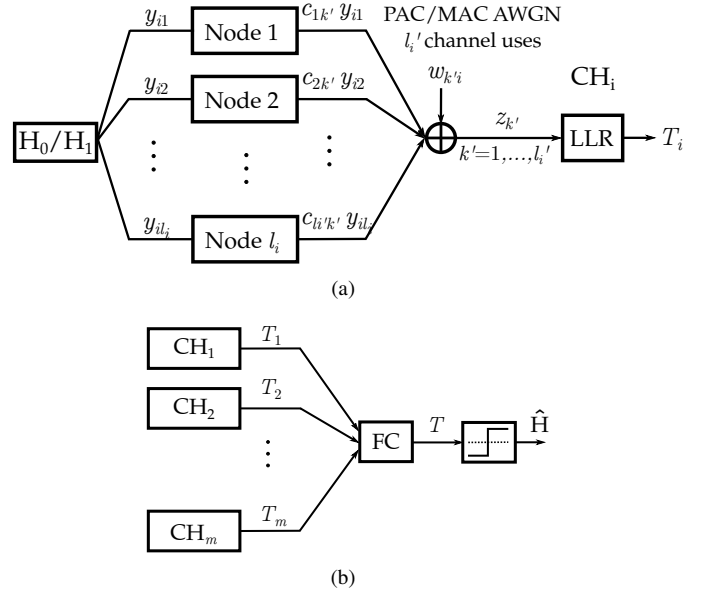


Fig. 2. A distributed scheme for gathering information in a WSN. (a) The nodes use a PAC or a MAC channel to communicate with their corresponding CH. (b) The FC collects the processed information of each CH and takes a decision.

#### A. The PAC precoding strategy

If a PAC is considered, the received signal at the  $i$ -th CH is

$$z_k = a_k c_k y_k + w_k, \quad k \in \mathcal{U}_i, \quad (3)$$

<sup>1</sup>We say that a channel use takes place when a symbol is transmitted.

where  $w_k$  is the zero-mean communication white noise with variance  $\sigma_w^2$  and circularly-symmetric complex Gaussian distribution independent of the measurements  $y_k$ ,  $a_k = \|\mathbf{x}_k - \mathbf{x}_{\text{CH}_i}\|^{-\frac{\epsilon}{2}}$  is the factor associated with the path loss effect between the position of the node  $\mathbf{x}_k \in \mathbb{R}^2$  and the position of the corresponding cluster head  $\mathbf{x}_{\text{CH}_i} \in \mathbb{R}^2$ ,  $\epsilon$  is the path loss exponent,  $c_k$  is the precoding coefficient used by the  $k$ -th sensor, and  $\mathcal{U}_i \subseteq \mathcal{C}_i$  is the set of channel uses indexes for the  $i$ -th cluster with cardinality  $|\mathcal{U}_i| = l'_i$ . Under this strategy, each sensor has a single channel use to communicate its measurement. If only  $l'_i < l_i$  channel uses are available,  $l_i - l'_i$  sensors in the cluster stay silent. Let  $C_i = \text{diag}((c_k)_{k \in \mathcal{U}_i})$  be the *precoding matrix* of the  $i$ -th cluster,  $A_i = \text{diag}((a_k)_{k \in \mathcal{U}_i})$ , and define the column vectors  $\mathbf{z}_i = [z_k]_{k \in \mathcal{U}_i}$ , and  $\mathbf{w}_i = [w_k]_{k \in \mathcal{C}_i'}$  with covariance matrix  $\sigma_w^2 I_{l'_i}$ . The received signal vector in each cluster can be expressed as

$$\mathbf{z}_i = A_i C_i \mathbf{y}_i + \mathbf{w}_i, \quad i = 1, \dots, m. \quad (4)$$

with covariance matrix under  $\mathcal{H}_j$ ,  $j = 0, 1$  given by,

$$\Xi_{j,l_i} = C_i^H A_i \Sigma_{j,l_i} A_i C_i + \sigma_w^2 I_{l'_i}. \quad (5)$$

The average energy consumed by the  $k$ -th node in the corresponding  $i$ -th cluster during one of the  $l'_i$  scheduled transmissions is:

$$E_k = \mathbb{E}(|c_k y_k|^2) = |c_k|^2 (\sigma_v^2 + p_1 \sigma_s^2), \quad (6)$$

where  $p_1$  is the *a priori* probability of the state of nature  $\mathcal{H}_1$ . When  $p_1$  is unknown, a natural upper-bound for  $E_k$  is obtained by taking  $p_1 = 1$ . In this work, we will consider the following individual sensor energy constraint:

$$E_k \leq \bar{E} \quad \forall k \in [1 : n]. \quad (7)$$

For the particular case when the sensed process is uncorrelated, and considering an individual sensor energy constraint, it can be easily proved that the optimal strategy consists in using all the available energy at each sensor. In this paper, we will consider this strategy even for correlated processes. The precoding strategy for a PAC channel named amplify and forward strategy (AFS-PAC) is defined next.

**Definition 1** (AFS-PAC). Considering the  $i$ -th cluster, (6) and (7), the precoding matrix for the AFS-PAC strategy is

$$C_i = \sqrt{\frac{\bar{E}}{\sigma_s^2 + \sigma_v^2}} I_{l'_i}, \quad i = 1, \dots, m. \quad (8)$$

### B. The MAC precoding strategies

In the case that a MAC is used, each sensor communicates with its corresponding CH using different gains for each channel use. The signal collected at the  $i$ -th CH is a noisy version of the coherent superposition of the symbols transmitted by the  $l_i$  sensors through the corresponding MAC:

$$z_{k'} = \sum_{k \in \mathcal{C}_i} a_k c_{kk'} y_k + w_{k'}, \quad k' \in \mathcal{U}_i, \quad (9)$$

where  $a_k$ ,  $w_{k'}$  and  $\mathcal{U}_i$  are the same quantities defined above. Thus, the  $k$ -th sensor transmits in the  $k'$ -th channel use

its measurement  $y_k$  scaled by the precoding coefficient  $c_{kk'}$  through the MAC channel. Stacking the measurements from each cluster, the same vector signal model (4) is obtained, although for the MAC channel, the precoding matrix defined by  $C_i = [c_{kk'}]_{k \in \mathcal{C}_i, k' \in \mathcal{U}_i}$  is not, in general, a diagonal matrix. Considering (9), the average energy consumed by the  $k$ -th node in the corresponding  $i$ -th cluster during the  $l'_i$  scheduled MAC channel uses is:

$$E_k = \mathbb{E} \left( \sum_{k' \in \mathcal{U}_i} |c_{kk'} y_k|^2 \right) = (\sigma_v^2 + p_1 \sigma_s^2) \sum_{k' \in \mathcal{U}_i} |c_{kk'}|^2. \quad (10)$$

If  $\{s_k\}$  is spatial-stationary, two schemes presented in [14] may be used to optimally allocate the energy and bandwidth in successive uses of the MAC channel. Both schemes, called *principal components strategy in a multiple-access channel* (PCS-MAC) and *principal frequency strategy in a multiple-access channel* (PFS-MAC), are asymptotically the best strategies among the orthogonal schemes considering energy and bandwidth constraints and the error exponents as metrics of performance. Although error exponents are sometimes useful to compare and define strategies of large networks, they do not provide an exact characterization of the error probabilities, which are the typically required metrics of performance for WSNs. Therefore, in this work, we mainly use the error probabilities as metrics of performance. Because the optimum strategy for a MAC, in terms of error probabilities, remains an open problem, we will use the PFS-MAC strategy (which is optimal in terms of error exponents) and show that good results are obtained.

On the other hand, if  $\{s_k\}$  is not spatial-stationary, in general, the covariance matrix of the process does not have a regular structure (i.e. Toeplitz form). However, in this case the PCS-MAC scheme presented in [14, Def. 2] may be used to optimally allocate the transmission energy and bandwidth. This scheme requires a built infrastructure that might not be available on any WSN. For that reason, in the spatial non-stationary case, we will consider the AFS-PAC strategy.

Next, we describe the PFS-MAC strategy for stationary processes. During each channel use, each CH receives a noisy version of a given frequency bin of the discrete Fourier transform (DFT) of the corresponding measurement vector  $\mathbf{y}_i$ . The idea behind this strategy is to transmit to the corresponding CH the *most distinguishing* spectral components of the power spectral density (PSD) of the sensed process  $\phi(\nu)$  (achieved when  $l$  grows unbounded). In particular, we will use for each cluster a simple almost-optimal solution that employs the PFS-MAC scheme with a constant energy profile (EP) on a given set of transmission frequencies. The strategy, called ON/OFF-EP (see [14]), is defined as follows:

**Definition 2** (PFS-MAC with ON/OFF-EP). Consider the  $i$ -th cluster. For a given  $l_i$ , denote by  $(j_1, j_2, \dots, j_{l'_i})$  a permutation of  $\{1, 2, \dots, l_i\}$  such that  $\phi\left(\frac{j_1}{l_i}\right) \geq \phi\left(\frac{j_2}{l_i}\right) \geq \dots \geq \phi\left(\frac{j_{l'_i}}{l_i}\right)$ . Let  $l'_i$  be the number of channel uses or bandwidth constraint.

Let  $\beta_i = \frac{l'_i}{l_i}$  be the fraction of channel uses (or *bandwidth*) used at cluster  $i$ . Then, the set of channel uses is defined by  $\mathcal{U}_i = \{j_1, \dots, j_{l'_i}\}$  and the precoding matrix of the PFS-MAC strategy with ON/OFF-EP is

$$C_i = \gamma_i A_i^{-1} F_{l_i l'_i}, \quad i = 1, \dots, m, \quad (11)$$

where  $F_{l_i l'_i} = [\zeta_{j_1}, \dots, \zeta_{j_{l'_i}}] \in \mathbb{C}^{l_i \times l'_i}$  is a sub-matrix of the normalized DFT matrix of size  $l_i \times l_i$ , i.e.,  $\zeta_{k'} = [f_{1k'}, f_{2k'}, \dots, f_{l_i k'}]^T$  with  $f_{kk'} = \frac{1}{\sqrt{l_i}} \exp\left(j2\pi \frac{(k-1)(k'-1)}{l_i}\right)$ ,  $k = 1, \dots, l_i$ ; the diagonal matrix  $A_i^{-1}$  performs a channel inversion to compensate for the different gains introduced by the path loss. Using (11) in (10), and considering the individual sensor energy constraint (7), we obtain that  $\gamma_i$  is limited by the longest distance between the nodes and the CH of the  $i$ -th cluster. Then,

$$\gamma_i = \sqrt{\frac{\bar{E}}{(\sigma_s^2 + \sigma_v^2) \beta_i (d_{\max, i})^\epsilon}}, \quad i = 1, \dots, m, \quad (12)$$

where  $d_{\max, i} = \max_{k \in C_i} \|\mathbf{x}_k - \mathbf{x}_{CH_i}\|$ .

Using the above scheme, the covariance matrices of  $\mathbf{z}_i$  under  $\mathcal{H}_1$  and  $\mathcal{H}_0$  are, respectively,

$$\Xi_{1, l'_i} = \gamma_i^2 F_{l_i l'_i}^H \Sigma_{s, l_i} F_{l_i l'_i} + (\gamma_i^2 \sigma_v^2 + \sigma_w^2) I_{l'_i} \quad (13)$$

$$\Xi_{0, l'_i} = (\gamma_i^2 \sigma_v^2 + \sigma_w^2) I_{l'_i}. \quad (14)$$

#### IV. DETECTION STATISTICS

##### A. Full Correlation Strategy

Suppose that the FC has direct access to all the measurements available in the CHs  $\mathbf{z} = [\mathbf{z}_1^T, \dots, \mathbf{z}_m^T]^T$  through ideal noiseless channels. The covariance matrix of  $\mathbf{z}$  under  $\mathcal{H}_j$ ,  $j = 0, 1$ , is

$$\Xi_{i, n'} = B_{nn'}^H \Sigma_{j, n} B_{nn'} + \sigma_w^2 I_{n'}, \quad (15)$$

where  $B_{nn'} = \text{diag}(A_1 C_1, \dots, A_m C_m)$ . Consider the Neyman-Pearson problem [25] for a fixed false alarm probability level  $\alpha$ , where a false alarm event occurs when  $\mathcal{H}_1$  is declared but  $\mathcal{H}_0$  is true. Let the logarithmic likelihood ratio (LLR) [25] of the full correlation strategy (FCS) be:

$$T_{\text{FCS}}(\mathbf{z}) = \log \frac{p(\mathbf{z}|\mathcal{H}_1)}{p(\mathbf{z}|\mathcal{H}_0)} \equiv \mathbf{z}^H \left( \Xi_{0, n'}^{-1} - \Xi_{1, n'}^{-1} \right) \mathbf{z}, \quad (16)$$

where we discarded the constant term  $\log \det(\Xi_{0, n'} \Xi_{1, n'}^{-1})$  without affecting the performance of the statistic. Now, under the Neyman-Pearson setting, the optimal decision rule chooses  $\mathcal{H}_1$  if  $T_{\text{FCS}}(\mathbf{z}) > \tau_n$ , and  $\mathcal{H}_0$  otherwise, where the threshold of the test  $\tau_n$  depends on  $\alpha$ .

In a distributed setting, with noisy links from the CHs to the FC, this statistic provides a lower bound for the false alarm and miss-detection error probabilities. It is worth to mention that these lower bounds are not necessarily tight due to the fact that they do not contemplate the degradation effect introduced by the communication channel between the CHs and the FC.

##### B. Local Correlation Strategy

When the process  $\{s_k\}$  is spatially correlated, the covariance matrix  $\Xi_{1, n'}$  in (15) is neither diagonal nor block-diagonal. Hence, (16) cannot be expressed as the sum of the local LLRs from each cluster, and processing the data independently does not lead to the global LLR. However, it is possible to implement simple distributed detection schemes to build an appropriate statistic at the FC using the compressed data transmitted by the CHs.

The FCS statistic can be decomposed in such a way that each term is a function of one or at most two sets of measurements from each cluster, i.e.,

$$T_{\text{FCS}}(\mathbf{z}) = \sum_{i=1}^m \mathbf{z}_i^H \left( \Xi_{0, l'_i}^{-1} - \Xi_{1, l'_i}^{-1} \right) \mathbf{z}_i + \sum_{i=1}^m \mathbf{z}_i^H M_{ii} \mathbf{z}_i + \sum_{i=1}^m \sum_{j=1, j \neq i}^m \mathbf{z}_i^H M_{ij} \mathbf{z}_j. \quad (17)$$

where  $\Xi_{0, l'_i}$  and  $\Xi_{1, l'_i}$  are the  $(i, i)$ -th block of  $\Xi_{0, n'}$  and  $\Xi_{1, n'}$ , respectively,  $M_{ii}$  is the difference between the  $(i, i)$ -th block of  $\Xi_{1, n'}^{-1}$  and  $\Xi_{1, l'_i}^{-1}$ ; and  $M_{ij}$  is the  $(i, j)$ -th block of  $\Xi_{1, n'}^{-1}$ . The first term in (17) captures most of the correlation between sensor measurements of the same cluster. We refer to this as the *intra-cluster* correlation term. The second term captures part of the intra-cluster correlation not included by the first term. The third term considers the correlation between sensor measurements of different clusters and we refer to this as the *inter-cluster* correlation term.

Considering that the correlation between measurements is a local effect, it is expected that the inter-cluster correlation term does not affect too much the whole statistic if each cluster has an appropriate size respect to the coherence distance of the process<sup>2</sup>. Keeping this effect in mind and considering also that gathering measurements from distant nodes is energy expensive, it would be reasonable to use a scheme that takes into account only the intra-cluster correlation. Then, the FC could process the data sent by the CHs without considering the inter-cluster correlation. The strategy is named local correlation strategy (LCS).

Let the local LLR at the  $i$ -th cluster be:

$$T_{\text{LCS}, i}(\mathbf{z}_i) = \log \frac{p(\mathbf{z}_i|\mathcal{H}_1)}{p(\mathbf{z}_i|\mathcal{H}_0)} \equiv \mathbf{z}_i^H \left( \Xi_{0, l'_i}^{-1} - \Xi_{1, l'_i}^{-1} \right) \mathbf{z}_i. \quad (18)$$

Each CH is able to exploit the intra-cluster correlation by computing the local LLR  $T_{\text{LCS}, i}(\mathbf{z}_i)$  while keeping the communication energy cost relatively low. Moreover, no additional overhead is necessary to communicate and/or estimate the ACF between measurements of different clusters. Only the local intra-cluster ACF need to be known or estimated at each CH. In the next step of the scheme, each CH transmits its corresponding  $T_{\text{LCS}, i}(\mathbf{z}_i)$  to the FC. This communication can be done using a PAC or a MAC. When a PAC is used, since the inter-cluster correlation is discarded, the FC builds the global

<sup>2</sup>Measurements taken in sensors separated by a distance greater than the coherence distance of the process are weakly correlated.

statistic adding the local statistics provided by the CHs, as suggested by the first term of (17):

$$T_{\text{LCS}}(\mathbf{z}) = \sum_{i=1}^m \mathbf{z}_i^H \left( \Xi_{0,l'_i}^{-1} - \Xi_{1,l'_i}^{-1} \right) \mathbf{z}_i. \quad (19)$$

On the other hand, if a MAC is used, the same statistic can be obtained. Notice, however, that the sum of the local statistics is naturally performed by sending them synchronously through the MAC channel. The channel communication between the CHs and the FC is assumed to be noiseless. We can justify this saying that usually there are much less CHs than nodes in the network and they typically have some infrastructure that the nodes do not have. For example, the CHs could have high directive antennas pointed to the FC, better energy budgets than the nodes, and, for the strategies considered in this paper, the CHs transmit scalar numbers which could be communicated using a reliable low rate coded digital modulation. For later developments, it is convenient to express the statistic (19) as follows

$$T_{\text{LCS}}(\mathbf{z}) = \mathbf{z}^H \Gamma_{n'} \mathbf{z}. \quad (20)$$

where  $\Gamma_{n'} = \text{diag} \left( \Xi_{0,l'_1}^{-1} - \Xi_{1,l'_1}^{-1}, \dots, \Xi_{0,l'_m}^{-1} - \Xi_{1,l'_m}^{-1} \right)$  is a  $n' \times n'$  block diagonal matrix, with  $n' = \sum_{i=1}^m l'_i$ . Notice that, in general,  $\Gamma_{n'} \neq \Xi_{0,n'}^{-1} - \Xi_{1,n'}^{-1}$ . Hence, the first term of (20) is not exactly the LLR of  $\mathbf{z}$  although it could be a good approximation if the size of the clusters are greater than the coherence distance of the correlated process. In the FC, the statistic  $T_{\text{LCS}}(\mathbf{z})$  is compared against a threshold  $\tau_n$  in order to decide about the true state of nature.

Now, we can elaborate an analytical justification about why we ignore the inter-cluster correlation in the LCS strategy using (17). The difference between the FCS and LCS statistics is the matrix used to produce the quadratic form: while  $\Xi_{0,n'}^{-1}$  is the same in both statistics because it is a diagonal matrix,  $\Xi_{1,n'}^{-1} \neq \text{diag}(\Xi_{1,l'_1}^{-1}, \dots, \Xi_{1,l'_m}^{-1})$ . However, it can be shown, by using the Schur complement, that the Taylor approximation of order 0 of the  $l'_i \times l'_i$  block matrices of  $\Xi_{1,n'}^{-1}$  is exactly the block diagonal matrix  $\text{diag}(\Xi_{1,l'_1}^{-1}, \dots, \Xi_{1,l'_m}^{-1})$ . Clearly, this would be a good approximation when the inter-cluster correlation is weak.

### C. Energy detector

We also consider the statistic usually called energy detector, which does not contemplate the correlation of the sensed process. Therefore, it will result useful to compare it with the previous statistics to show how much one can loose if the correlation is not taken into account. Once each CH has the vector of measurements  $\mathbf{z}_i$  in (4), it builds  $T_{\text{ED},i}(\mathbf{z}_i) = \|\mathbf{z}_i\|^2$  and sends it through a PAC or a MAC to the FC. The energy detector available in the FC is defined by

$$T_{\text{ED}}(\mathbf{z}) = \|\mathbf{z}\|^2. \quad (21)$$

Notice that this statistic is optimal only when the process  $\{s_k\}$  is white noise process.

### D. Clustering partition problem

Now, we are ready to formulate the main problem of this paper. We first summarize the whole strategy used. The  $n$  sensors of the WSN take measurements and transmit them to their corresponding CHs using the PFS-MAC strategy or the AFS-PAC strategy. Each CH has access to (4) and allows the FC to build a statistic  $T_j$  using one of the previous strategies:  $j \in \{\text{FCS}, \text{LCS}, \text{ED}\}$ . The FC makes a decision about the state of the nature comparing  $T_j$  against threshold  $\tau_j$ . The false alarm and the miss-detection probability are, respectively,  $P_{\text{fa}}^n = \mathbb{P}_0(T_j > \tau_j)$  and  $P_{\text{m}}^n = \mathbb{P}_1(T_j < \tau_j)$ .

The way each sensor node is assigned to each cluster is an important issue affecting the performance of the network. One may wonder which is the best allocation possible and which is its level of performance. To that end we can fix the probability of false alarm to a level  $\alpha \in (0, 1)$  and consider an energy constraint given by (7). Let  $e_k$  be the label of the  $k$ -th sensor that identifies to which cluster it belongs, i.e.,  $e_k = i$  if  $k$  belongs to the  $i$ -th cluster. We can then formulate a problem in order to obtain the optimal clustering partition that could be cast as follows:

$$\begin{aligned} & \inf_{(e_1, \dots, e_n) \in [1:m]^n} P_{\text{m}}^n, \\ \text{s.t. } & P_{\text{fa}}^n \leq \alpha, E_k \leq \bar{E}, \forall k \in [1:n]. \end{aligned} \quad (22)$$

where the threshold of the test  $\tau_j$  depends on  $\alpha$ . This is a challenging non-linear integer programming problem with constraints. In fact, it is a NP-hard problem and its solution is out of the scope of this paper. As our objective is to understand the trade-off between the beneficial use of spatial correlation against energy consumption in order to overcome the path-loss effect, we will consider a simpler clustering partition problem. We will assume that all the clusters have the same shape, in particular, we will assume that each of the  $m$  clusters are square regions of area  $l_i A_0$ , where  $l_i$  is the number of sensors in the  $i$ -th cluster, and  $n A_0 = \sum_{i=1}^m l_i A_0$  is total area in which the network is deployed. This simple model captures the basic fact that when the number of clusters is large, the size of them will be small, and the path loss effect in the local transmissions will be less harmful. However, as the clusters are smaller, less sensor nodes are present within them. In addition, some other nodes in their local neighborhoods, which can have strongly correlated measurements, could be assigned to other adjacent clusters. As the clusters do not cooperate and they make their processing independently, this clearly reduces the benefit of using the spatial correlation in each cluster. This basic model will be used in the following section with minor variations. For example, in Section VI-A we will assume that each cluster has exactly  $l$  nodes arbitrary distributed in space. On the other hand, in Section VI-B, we will assume that the nodes are spatially distributed as a homogeneous Poisson Point Process (PPP) [26] where the average number of sensor per cluster is  $l$ , although for each realization of the spatial random process each cluster could have different number of nodes. In both cases, we variate the area of each cluster and, consequently, the number of sensors per cluster  $l$  (or its average) but keeping

constant the total amount of nodes  $n$  in the sensor network (or its average).

## V. COMPUTATION OF THE ERROR PROBABILITIES

In this section, we present a result used to compute tight approximations of the miss-detection probability  $P_m^n$  and the false alarm probability  $P_{fa}^n$  using the LMGFs of the statistic under  $\mathcal{H}_0$  and  $\mathcal{H}_1$ .

**Theorem 1.** Let  $\{y_k\}_{k=1}^n$  be mutually independent random variables with probability density function (PDF)  $p_k$  and LMGF  $\mu_k(s) \triangleq \log \mathbb{E}(e^{y_k s})$ . Assume that  $\mathbb{E}(y_k^2)$  and  $\mathbb{E}(|y_k - \mathbb{E}(y_k)|^3)$  exist and are finite  $\forall k \in [1 : n]$ . Let  $T_n = y_1 + \dots + y_n$  with LMGF  $\mu_{T_n} \triangleq \log \mathbb{E}(e^{T_n s})$  and let  $\tau_n \in \mathbb{R}$ . Then, if  $\tau_n > \mathbb{E}(T_n)$ ,

$$\mathbb{P}(T_n > \tau_n) = \left( \frac{1}{\sqrt{2\pi s_0^2 \dot{\mu}_{T_n}(s_0)}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right) e^{-(s_0 \dot{\mu}_{T_n}(s_0) - \mu_{T_n}(s_0))} \quad (23)$$

where  $s_0 > 0$  is the solution to  $\tau_n = \dot{\mu}_{T_n}(s_0)$ . On the other hand, if  $\tau_n < \mathbb{E}(T_n)$ ,

$$\mathbb{P}(T_n < \tau_n) = \left( \frac{1}{\sqrt{2\pi s_1^2 \dot{\mu}_{T_n}(s_1)}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right) e^{-(s_1 \dot{\mu}_{T_n}(s_1) - \mu_{T_n}(s_1))} \quad (24)$$

where  $s_1 < 0$  is the solution to  $\tau_n = \dot{\mu}_{T_n}(s_1)$ .

*Proof:* See App. I-A. ■

**Corollary 1** (Gaussian quadratic form). Consider the statistic given by the quadratic form

$$T_{n'} = \mathbf{z}^H P_{n'} \mathbf{z}, \quad (25)$$

where the distribution of  $\mathbf{z}$  under  $\mathcal{H}_i$  is  $\mathcal{CN}(\mathbf{0}, \Xi_{i,n'})$ ,  $i = 0, 1$ . Assume that  $\Xi_{i,n'}$  and  $P_{n'}$  are positive definite matrices and that  $\tau_{n'} \in (\mathbb{E}_0(T_{n'}), \mathbb{E}_1(T_{n'}))$  is the threshold of the test. The false alarm and the miss error probabilities are, respectively,

$$\begin{aligned} P_{fa}^{n'} &= \left( \frac{1}{\sqrt{2\pi s_0^2 \text{tr}\{[(\Xi_{0,n'} P_{n'})^{-1} - s_0 I_{n'}]^{-2}\}}} + \mathcal{O}\left(\frac{1}{\sqrt{n'}}\right) \right) \times \\ &\quad \times e^{-s_0 \text{tr}\{[(\Xi_{0,n'} P_{n'})^{-1} - s_0 I_{n'}]^{-1}\} - \log \det(I_{n'} - s_0 \Xi_{0,n'} P_{n'})}, \\ P_m^{n'} &= \left( \frac{1}{\sqrt{2\pi s_1^2 \text{tr}\{[(\Xi_{1,n'} P_{n'})^{-1} - s_1 I_{n'}]^{-2}\}}} + \mathcal{O}\left(\frac{1}{\sqrt{n'}}\right) \right) \times \\ &\quad \times e^{-s_1 \text{tr}\{[(\Xi_{1,n'} P_{n'})^{-1} - s_1 I_{n'}]^{-1}\} - \log \det(I_{n'} - s_1 \Xi_{1,n'} P_{n'})}, \end{aligned}$$

where  $s_0 > 0$  and  $s_1 < 0$  are the solution to  $\tau_{n'} = \text{tr}\{[(\Xi_{0,n'} P_{n'})^{-1} - s_0 I_{n'}]^{-1}\}$  and  $\tau_{n'} = \text{tr}\{[(\Xi_{1,n'} P_{n'})^{-1} - s_1 I_{n'}]^{-1}\}$ , respectively, and make  $I_{n'} - s_i \Xi_{i,n'} P_{n'}$  positive definite for  $i = 0, 1$ .

*Proof:* See App. I-B. ■

## VI. PERFORMANCE ANALYSIS

In Section VI-A, we consider a case of study where the process  $\{s_k\}$  is assumed to be equicorrelated, a particular case of a stationary process. It properly models the spatial correlation between the nodes and allows to obtain closed

forms expressions of the error probabilities. In Section VI-B, we will consider that the nodes are spatially distributed as a PPP. We will see that similar conclusions to those obtained for the equicorrelated process can be drawn for non-stationary process. Before continuing, let  $\text{SNR}_M = \frac{\sigma_s^2}{\sigma_w^2}$  and  $\text{SNR}_C = \frac{\bar{E}}{\sigma_w^2}$  be the signal to noise ratio of the measurements and of the communication channel, respectively.

### A. A case of study: the equicorrelated process

Consider a stationary zero-mean circular-symmetric Gaussian equicorrelated process with ACF  $R_s(k) = \sigma_s^2(1-\rho)\delta_{k0} + \sigma_s^2\rho$ , where  $\delta_{k0}$  is the delta of Kronecker. The PSD of the equicorrelated process is  $\phi(\nu) = \sigma_s^2(1-\rho) + \sigma_s^2\rho\delta(\nu)$ , where  $\delta(\nu)$  is the Dirac delta distribution and  $\rho$  is a scalar parameter that controls the correlation,  $0 \leq \rho < 1$ . The covariance matrix for the measurements of each cluster is given by  $\Sigma_{s,l} = \sigma_s^2(1-\rho)I_l + \sigma_s^2\rho\mathbf{1}_l$  where  $\mathbf{1}_l$  is a  $l \times l$  matrix of ones.

As the equicorrelated process is stationary, both PFS-MAC or AFS-PAC strategies can be used. However, in this section we will focus on the first one. We assume that all the clusters are identical and have the same parameters in Def 2:  $l_i = l$ ,  $l'_i = l'$ ,  $\beta_i = \beta$ ,  $d_i = d_{\max}(l)$  and  $\gamma_i = \gamma_0 \forall i$ . Considering that the CH could be in any position inside the coverage area  $lA_0$  of its corresponding cluster, and that each cluster is a square region, the maximum possible distance between a node and its corresponding CH is  $d_{\max}(l) = \sqrt{2lA_0}$ .

Let  $\text{SNR}_{\text{PFS-MAC}}(l) = \frac{\gamma_0^2 \sigma_s^2}{\gamma_0^2 \sigma_s^2 + \sigma_w^2}$  be the effective signal-to-noise ratio of the measurements available in each CH. If we replace (12) in the previous definition, we have

$$\text{SNR}_{\text{PFS-MAC}}(l) = \frac{\text{SNR}_M \text{SNR}_C}{\text{SNR}_C + (1 + \text{SNR}_M)(2lA_0)^{\frac{\epsilon}{2}}}, \quad (26)$$

where we have emphasized that the effective signal-to-noise ratio is a function of  $l$ .

In order to evaluate the error probabilities of the LCS, FCS, and ED schemes, we need to compute the LMGFs of each statistic under  $\mathcal{H}_i$  and their first and second derivatives to then apply Th. 1. For this particular process, it is possible to obtain closed forms solutions for the error probabilities for any threshold  $\tau_n$ . However, these expressions are lengthy and are not shown here. The symbolic expressions are shown in the MATLAB script *SymExprForPfaPm.m* available with this paper.

Therefore, we make focus on the case when the threshold for each strategy is set such that the miss error exponent is zero, where simple expressions are obtained. We summarize the results in the following lemma.

**Lemma 1** (Probabilities of false alarm for the equicorrelated process using the PFS-MAC strategy). Consider the following thresholds such that the miss error exponents are zero for the three strategies:  $\tau_{\text{LCS}} = \mathbb{E}_1(T_{\text{LCS}}(\mathbf{z}))$ ,  $\tau_{\text{FCS}} = \mathbb{E}_1(T_{\text{FCS}}(\mathbf{z}))$  and  $\tau_{\text{ED}} = \mathbb{E}_1(T_{\text{ED}}(\mathbf{z}))$ . Then, the false alarm probabilities of the

LCS, FCS and ED strategies using PFS-MAC are, respectively,

$$P_{fa,LCS}^n = \left( \frac{1}{\Gamma \sqrt{2\pi n[(l+\beta-2)\rho^2+2(1-\beta)\rho+\beta]}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right) \times e^{-n(\Gamma(\rho+\beta(1-\rho))-\frac{1}{l}\log(1+\Gamma(1+(l-1)\rho))-(\beta-\frac{1}{l})\log(1+\Gamma(1-\rho)))} \quad (27)$$

$$P_{fa,FCS}^n = \left( \frac{1}{\Gamma \sqrt{2\pi n[(n+\beta-2)\rho^2+2(1-\beta)\Gamma+\beta]}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right) \times e^{-n(\Gamma(\rho+\beta(1-\rho))-\frac{1}{n}\log(1+\Gamma(1+(n-1)\rho))-(\beta-\frac{1}{n})\log(1+\Gamma(1-\rho)))} \quad (28)$$

$$P_{fa,ED}^n = \left( \frac{1}{\Gamma \sqrt{2\pi n\beta}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right) e^{-n\beta(\Gamma-\log(1+\Gamma))}, \quad (29)$$

where  $\Gamma \triangleq \text{SNR}_{\text{PFS-MAC}}(l)$ .

*Proof:* See App. II

It is easy to show that  $P_{fa,LCS}^n$  decreases monotonically with  $n$ ,  $\beta$  and  $\rho$ . If  $\text{SNR}_{\text{PFS-MAC}}(l)$  were independent of  $l$ ,  $P_{fa,LCS}^n$  would also decrease monotonically with  $l$  and  $\text{SNR}_{\text{PFS-MAC}}(l)$ . However, as it is seen in (26),  $\text{SNR}_{\text{PFS-MAC}}(l)$  decreases with  $l$ , and therefore, a trade-off between these two parameters is established: for a given spatial density of sensors, to increase  $l$ , the number of nodes on each cluster, permits to successfully exploit the correlation, although this produces a worse  $\text{SNR}_{\text{PFS-MAC}}(l)$  at each cluster head. If the cluster size is reduced,  $\text{SNR}_{\text{PFS-MAC}}(l)$  improves but the correlation is not fully exploited.

In Fig. 3, the false alarm probability of the three schemes LCS, FCS and ED given by (27), (28) and (29) respectively, are plotted against  $l$ , for the parameters shown in the caption. We observe that for the LCS scheme, to increase the size of

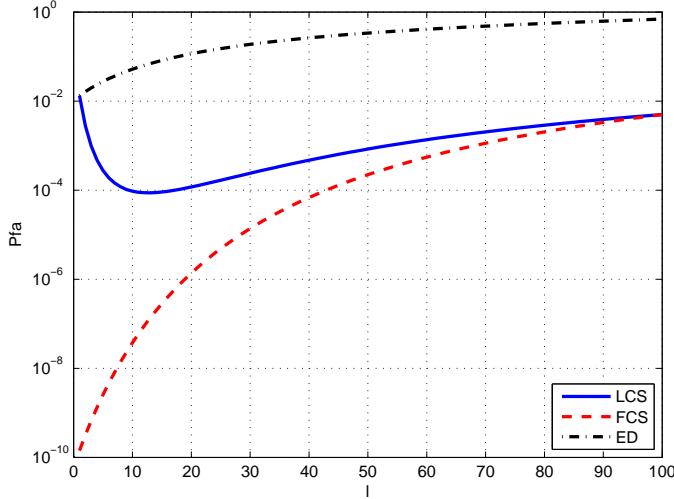


Fig. 3. Trade-off between exploiting correlation and improving the measurement SNR. Parameters:  $\text{SNR}_C = 12$  dB,  $\text{SNR}_M = -6$  dB,  $\rho = 0.9$ ,  $\epsilon = 2$ ,  $\beta = 1$  and  $n = 100$ .

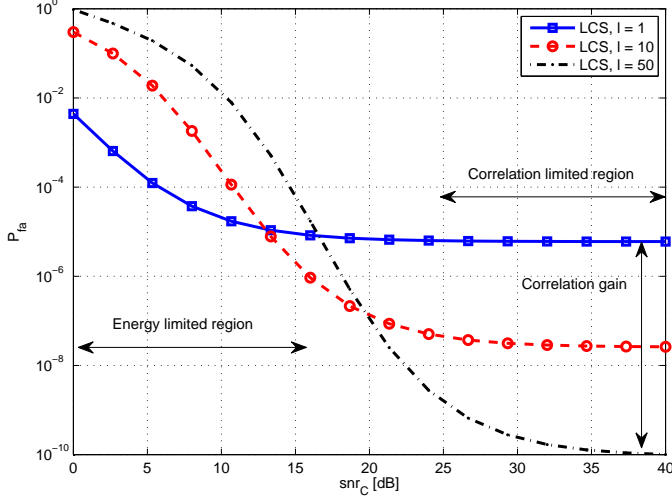
the cluster allows to exploit the correlation of the process and decrease  $P_{fa,LCS}^n$ . Beyond a certain cluster size, say  $l = 10$ , the benefit of exploiting the correlation is overcome by the loss of  $\text{SNR}_{\text{PFS-MAC}}(l)$ , therefore, causing a worse performance. In the case of ED, increasing the cluster size reduces the

effective  $\text{SNR}_{\text{PFS-MAC}}(l)$  without exploiting the correlation, due to the nature of this scheme. On the other hand, as the FCS scheme has access to the whole vector of measurements  $\mathbf{z}$ , and computes the optimum LLR detector, it is only limited by  $\text{SNR}_{\text{PFS-MAC}}(l)$ , which decreases when  $l$  increases. For a fixed cluster size, the FCS performance allows to quantify how much one can gain potentially by exploiting the correlation of the sensed process. For example, if  $l = 10$ , one could expect to decrease  $P_{fa,LCS}^n$  between 3 and 4 orders of magnitude, although this is not entirely true given that the FCS scheme permits the FC to have access to the full vector  $\mathbf{z}$  neither being penalized in bandwidth nor in transmitted energy, i.e., it could be a loose lower bound for some  $l$ 's. Additionally, in the case of the PFS-MAC scheme, the energy received in the CHs is limited by the furthest node, which could impose severe energy limitations. Notice that by definition of the LCS strategy, it coincides with the ED strategy when  $l = 1$  (no correlation is considered in the statistic) and with FCS strategy when  $l = n$  (the network has a unique cluster).

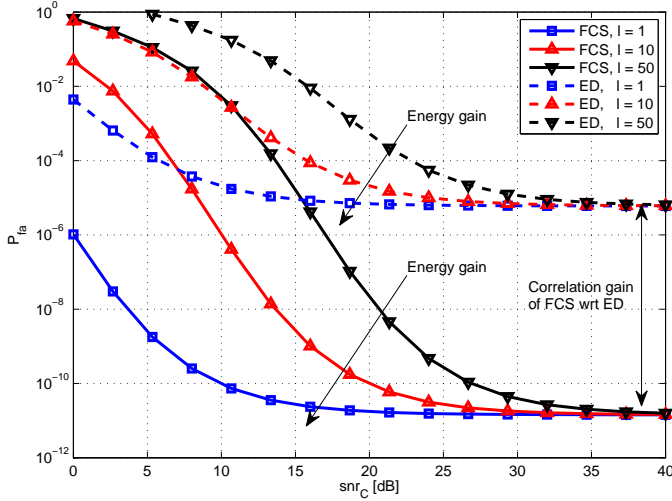
In Fig. 4, we plot the false alarm probability of the three statistics in Lem. 1 against  $\text{SNR}_C$  with  $l$  as a parameter. In Fig. 4(a) we plot only the LCS strategy for the sake of clarity. Two regimes can be observed there: i) an energy-limited regime for which  $\text{SNR}_C$  limits the performance and, ii) a correlation-limited regime, where the performance is limited due to fact that the inter-cluster correlation is not considered. When the cluster size is small, e.g. low number of sensors per cluster  $l$ , the sensors are near to the CH and the communication between them is reliable. We see in Fig. 4(a) that a WSN with small clusters performs better when the sensor energy budget is tight or the path loss effect is severe. However, when the sensor energy budget and the path loss effect are not an issue, large clusters are preferred because they can exploit successfully the intra-cluster correlation, obtaining large correlation gains even for weakly correlated processes ( $\rho = 0.2$ ), as indicated in the figure. In Fig. 4(b), we show the behavior of the FCS and the ED strategies. We see that both strategies have lesser false alarm probabilities for small cluster sizes for the entire  $\text{SNR}_C$  range. In other words, their performance is better due to an energy gain of the measurements in the CHs. Neither FCS nor ED modify how they consider the correlation of the process when the cluster size changes: the ED strategy always obviates the correlation while the FCS always considers the full correlation of the process (intra and inter-cluster). On the other hand, the proposed LCS makes the compromise between reliable measurement transmission to the CHs and exploiting the correlation of the process. For a given  $l$ , the FCS and the ED strategies provide a lower bound and an upper bound, respectively, of the error probability of the LCS strategy.

Similarly to Lem. 1, we use Th. 1 to compute the miss-detection probability for the three statistics LCS, FCS and ED, when the threshold for each test  $\tau_n$  is such that the false alarm error exponent is zero, i.e.,  $\tau_n$  is the mean of each statistic under  $\mathcal{H}_1$ . The expressions are lengthy and are not shown here. We have plotted them in Fig. 5, against  $\text{SNR}_C$  with  $l$  as a parameter. It is observed that, similarly to the probability of





(a) LCS strategy.

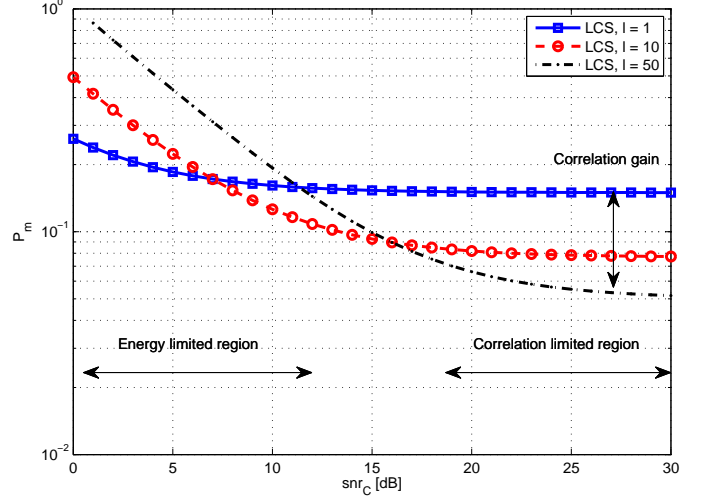


(b) FCS and ED strategies.

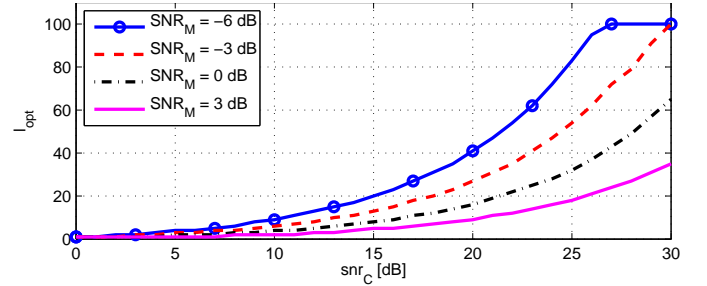
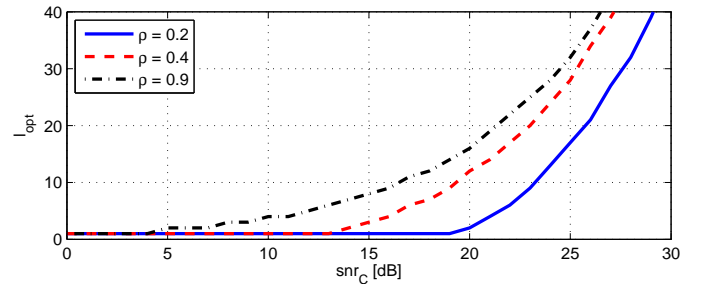
Fig. 4. Probability of false alarm versus communication signal to noise ratio with parameters  $\text{SNR}_M = -3$  dB,  $\rho = 0.2$ ,  $\epsilon = 2$ ,  $\beta = 1$  and  $n = 100$ .

false alarm, two regimes are observed for the miss-detection probability: a energy-limited regime and an correlation-limited regime. For a low  $\text{SNR}_C$ , the miss-detection probability is limited by  $\text{SNR}_{\text{PFS-MAC}}(l)$ . Therefore, a lower  $P_m$  is obtained when the cluster size is small (low  $l$ ). For a high  $\text{SNR}_C$ , the performance is limited by how much the correlation is exploited. As  $\text{SNR}_{\text{PFS-MAC}}(l)$  is not an issue, the LCS can increase the cluster size in order to exploit the intra-cluster correlation and decrease  $P_m$ .

To conclude this section, we obtain the optimum size  $l_{\text{opt}}$  which minimizes the false-alarm error probability of the LCS strategy given by (27) among all cluster sizes  $l \in [1 : n]$  such that  $n = ml$ . This is done numerically given that there is not a closed-form expression for  $l_{\text{opt}}$ . In Fig. 6(a), we plot  $l_{\text{opt}}$  versus  $\text{SNR}_C$  with  $\text{SNR}_M$  as a parameter. We see that the optimum cluster size increases with  $\text{SNR}_C$  for each  $\text{SNR}_M$ . We also observe that for a fixed  $\text{SNR}_C$ ,  $l_{\text{opt}}$  is greater for lower  $\text{SNR}_M$ . That is, as  $\text{SNR}_M$  decreases, it is better to exploit the intra-

Fig. 5.  $P_m$  versus  $\text{SNR}_C$ .  $\text{SNR}_M = 0$  dB,  $n = 100$  and  $\rho = 0.9$ .

cluster correlation increasing  $l$  than to increase  $\text{SNR}_{\text{PFS-MAC}}(l)$ . This is an important guideline for the case in which the WSN is composed by cheap nodes with limited sensing capabilities. In Fig. 6(b), we plot  $l_{\text{opt}}$  versus  $\text{SNR}_C$  with  $\rho$  as a parameter. For a given  $\text{SNR}_C$ , we see that it is more profitable to increase the size of the cluster when the process is highly correlated than to increase  $\text{SNR}_{\text{PFS-MAC}}(l)$ .

(a)  $\text{SNR}_M$  as a parameter.(b)  $\rho$  as a parameter.Fig. 6.  $l_{\text{opt}}$  versus  $\text{SNR}_C$  with  $\epsilon = 2$ ,  $\beta = 1$  and  $n = 100$ .

The optimum cluster size  $l_{\text{opt}}$  allows to redefine the idea of spatial coherence regions (SCRs). Typically, a SCR is an inherent property of the process to be sensed, uniquely determined by the spatial correlation between its samples [6]. In this work, guided by the error probability as an indicator of the ultimate performance, we obtain, that a SCR is not only

defined by the spatial correlation of the process, but also by critical parameters as  $\text{SNR}_M$ , transmitted energy  $\bar{E}$ , path loss exponent  $\epsilon$ , and bandwidth  $\beta$  available at each cluster.

### B. Random Network

In this section, we compute the performance of the statistics when the sensors of the network are spatially distributed as a homogeneous Poisson Point Process [26] of intensity  $\lambda > 0$ . In this case, the total number of sensors and the number of sensors per cluster follow a Poisson distribution with mean  $\lambda n A_0$  and  $\lambda A_0$ , respectively. For simplicity, we continue using the assumption that the clusters are squares of area  $l A_0$  and the region to be sensed has a total area  $n A_0$ . We set  $\lambda = 1$  and  $A_0 = 1$ . Therefore, the average number of sensors per cluster coincides with  $l$ . We also assume that the spatial process is isotropic meaning that the ACF depends only on the distance  $d$  between the points for which it is evaluated. Some examples of possible isotropic ACFs are the equicorrelated process presented previously, the exponential ACF  $R_s(d) = \sigma_s^2 \rho^d$  or the hyperbolic ACF  $R_s(d) = \sigma_s^2 \frac{\rho}{\rho + d}$ . The exponential ACF was found to be a good model for the shadowing effect in a urban cellular environment [28] and it will be used here.

The precoding strategy used for an arbitrary spatial distribution of the nodes is AFS-PAC whose gains are defined in (8). For this strategy, the signal-to-noise ratio for the  $k$ -th node at the  $i$ -th CH is defined as  $\text{SNR}_{\text{AFS-PAC}}(k) = \frac{c_{kk}^2 \sigma_s^2}{c_{kk}^2 \sigma_v^2 + \sigma_w^2 / a_k^2}$ . Replacing (8) in it gives

$$\text{SNR}_{\text{AFS-PAC}}(k) = \frac{\text{SNR}_C \text{SNR}_M}{\text{SNR}_C + (1 + \text{SNR}_M) \|\mathbf{x}_k - \mathbf{x}_{\text{CH},i}\|^\epsilon}.$$

We compute the average miss error probability  $\bar{P}_m$  as a metric of performance, defined by the average over the PPP  $\Phi$ ,  $\bar{P}_m = \mathbb{E}(P_m(\Phi))$  subject to  $P_{\text{fa}}(\Phi) = \alpha$  for each realization of the PPP. This average is computed using the method of Monte Carlo and the theoretical expressions found in Corollary 1 (labeled LCS, FCS and ED in Figs. 7 and 8). On the other hand, we also generate the sensed process and the corresponding noises using again the method of Monte Carlo to validate the theoretical expressions derived in Corollary 1 (labeled LCS-MC, FCS-MC and ED-MC in Figs. 7 and 8).

In Fig. 7 we plot the miss probability as a function of the average number of nodes per cluster  $l$  with the parameters indicated in its caption. Notice that the results for a random network with exponential ACF are similar to the equicorrelated process case. Considering the LCS strategy, in spite of decreasing the effective signal-to-noise ratio at each CH  $\text{SNR}_{\text{AFS-PAC}}(k)$  for most of the sensors when  $l$  increases up to 10, the miss error probability decreases because of the intra-cluster correlation is better exploited. If the cluster size increases further, the path loss effect inside each cluster dominates and the performance is deteriorated. The LCS strategy is close to the bound provided by the FCS scheme and the gain compared with the ED is considerable.

In Fig. 8 we plot the miss error probability as a function of  $\text{SNR}_C$  with the parameters indicated in its caption. Notice

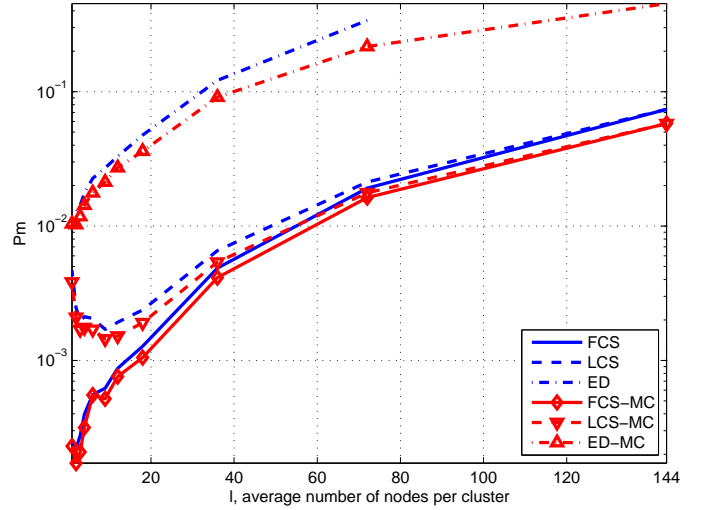


Fig. 7.  $\bar{P}_m$  versus  $l$  for an exponential ACF with  $\rho = 10$ ,  $n = 144$ ,  $P_{\text{fa}} = 10^{-2}$ ,  $\text{SNR}_M = 0$  dB and  $\text{SNR}_C = 12$  dB.

again that the results are similar to the closed-form expressions obtained for the equicorrelated process. The LCS strategy is close to the bound provided by the FCS scheme and the gain respect to the ED strategy is considerable for the entire shown range of  $\text{SNR}_C$ . These Monte Carlo curves also validate the results of Th. 1, where the central limit theorem was applied to compute the error probabilities, obtaining good approximations even for a low number of sensors  $n$ .

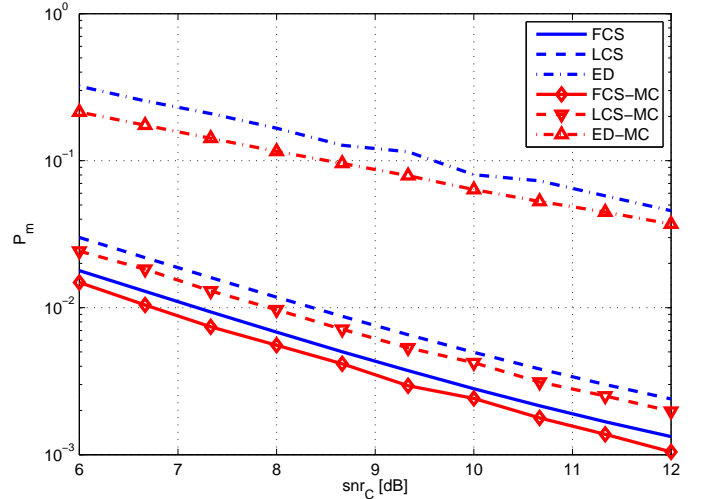


Fig. 8.  $P_m^n$  vs  $\text{SNR}_C$  for an exponential ACF with  $\rho = 10$ ,  $n = 144$ ,  $l = 18$ ,  $m = 8$ ,  $P_{\text{fa}} = 10^{-2}$ , and  $\text{SNR}_M = 0$  dB.

## VII. CONCLUSIONS

In this paper, we developed a model and appropriate tools for evaluating the performance of a distributed detection scheme that makes a trade-off between transmission energy and exploitation of the spatial correlation of the process. We characterized this trade-off and found two regimes of operation: an energy limited regime, and a correlation-limited

regime. We showed that the proposed LCS strategy allows to obtain excellent performances, which are close to the ones obtained with the use of the global network-wide correlation information. The use of the spatial correlation is of great importance in WSNs where the captured measurements have low signal-to-noise ratio and also when the process to be sensed presents moderate to high correlation.

## APPENDIX I ERROR PROBABILITY APPROXIMATIONS

In this appendix we derive an approximation for both the false alarm probability  $P_{fa}^n$  and the miss-detection probability  $P_m^n$  following a line reasoning similar to [29] and using a version of the central limit theorem (CLT) for a sum of independent but not identically distributed random variables given by the following lemma.

**Lemma 2** (Essen CLT). Let  $x_k$  be mutually independent random variables such that

$$\mathbb{E}(x_k) = 0, \quad \mathbb{E}(x_k^2) = \sigma_k^2, \quad \mathbb{E}(|x_k|^3) = \rho_k.$$

Let  $s_n^2 = \sigma_1^2 + \dots + \sigma_n^2$ ,  $r_n = \rho_1 + \dots + \rho_n$  and denote by  $F_n$  the cumulate distribution function (CDF) of  $(x_1 + \dots + x_n)/s_n$ . Then, for all  $u$  and  $n$ ,

$$|F_n(u) - \Phi(u)| \leq \frac{6r_n}{s_n^3},$$

where  $\Phi$  is the standard normal CDF.

*Proof:* See [30, p. 544, Th. 2]

### A. Proof of Theorem 1

We begin with the computation of  $\mathbb{P}(T_n > \tau_n)$  when  $\tau_n > \mathbb{E}(T_n)$ . Let  $q_k$  be the PDF of  $y_{k,s}$ , the tilted version of  $y_k$  with PDF  $p_k$ , such that  $q_k(y_{k,s}) = p_k(y_{k,s})e^{sy_{k,s} - \mu_k(s)}$ . It is straightforward to check that the first and second derivatives of  $\mu_k(s)$  satisfy  $\dot{\mu}_k(s) = \mathbb{E}(y_{k,s})$  and  $\ddot{\mu}_k(s) = \mathbb{V}\text{ar}(y_{k,s})$ , respectively. Define  $T_{n,s} = y_{1,s} + \dots + y_{n,s}$ . The PDF of  $T_{n,s}$  is

$$\begin{aligned} Q_T(t_s) &= \int_{\mathcal{R}} \prod_{k=1}^n q_k(y_{k,s}) dy_{k,s} \\ &= \int_{\mathcal{R}} \prod_{k=1}^n p_k(y_{k,s}) e^{sy_{k,s} - \mu_k(s)} dy_{k,s}, \end{aligned}$$

where the integration domain  $\mathcal{R}$  is defined as  $\mathcal{R} = \{(y_1, \dots, y_n) \in \mathbb{R}^n : y_{1,s} + \dots + y_{n,s} = t_s\}$ . Therefore,

$$\begin{aligned} Q_T(t_s) &= \left( \int_{\mathcal{R}} \prod_{k=1}^n p_k(y_{k,s}) dy_{k,s} \right) e^{st_s - \sum_{k=1}^n \mu_k(s)} \\ &= P_T(t_s) e^{st_s - \mu_T^n(s)}, \end{aligned}$$

where  $P_T$  is the PDF of  $T_n$  and  $\mu_T^n(s) = \sum_{k=1}^n \mu_k(s)$  is its LMGF. Thus,  $T_{n,s}$  is the tilted version of  $T_n$ . Then,

$$\begin{aligned} \mathbb{P}(T_n > \tau_n) &= \int_{\tau_n}^{\infty} P_T(t) dt \\ &= e^{-(s\tau_n - \mu_T^n(s))} \int_{\tau_n}^{\infty} Q_T(t) e^{-s(t - \tau_n)} dt \\ &= e^{-(s\dot{\mu}_T^n(s) - \mu_T^n(s))} \int_{\dot{\mu}_T^n(s)}^{\infty} Q_T(t) e^{-s(t - \dot{\mu}_T^n(s))} dt. \end{aligned} \quad (30)$$

(31)

The first factor in (30) is equal to the Chernoff bound if  $s$  is constrained to be positive. Because the LMGF  $\mu_T^n(s)$  is a convex function,  $\dot{\mu}_T^n(s)$  is a nondecreasing function. By assumption,  $\tau_n > \mathbb{E}(T_n)$ . Thus, the value of  $s$  that solves  $\tau_n = \dot{\mu}_T^n(s)$  satisfies  $s > 0$  and we can use this equality to replace  $\tau_n$  in (31).

The Chernoff error exponent defined as  $\lim_{n \rightarrow \infty} \frac{1}{n}(s\dot{\mu}_T^n(s) - \mu_T^n(s))$  is asymptotically tight, i.e., it is equal to the error exponent  $\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(T_n > \tau_n)$ . Therefore, we need to approximate the integral which is a sub-exponential factor. We let  $F_T$  be the CDF of  $Q_T$  and integrate by parts to obtain

$$\begin{aligned} \int_{\dot{\mu}_T^n(s)}^{\infty} Q_T(t) e^{-s(t - \dot{\mu}_T^n(s))} dt &= \\ &= \int_{\dot{\mu}_T^n(s)}^{\infty} s(F_T(t) - F_T(\dot{\mu}_T^n(s))) e^{-s(t - \dot{\mu}_T^n(s))} dt \\ &= s\sqrt{\ddot{\mu}_T^n(s)} \int_0^{\infty} (G_T(u) - G_T(0)) e^{-s\sqrt{\ddot{\mu}_T^n(s)}u} du, \end{aligned} \quad (32)$$

where in the second equality we have defined the following zero-mean unit-variance random variable:  $U_n = \frac{T_{n,s} - \dot{\mu}_T^n(s)}{\sqrt{\ddot{\mu}_T^n(s)}} = \frac{\sum_k y_{k,s} - \dot{\mu}_k(s)}{\sqrt{\ddot{\mu}_T^n(s)}}$  with CDF given by  $G_T$ . Applying Lem. 2 with  $x_k \triangleq y_{k,s} - \dot{\mu}_k(s)$ , it can be shown that  $s_n^2 = \ddot{\mu}_T^n(s) = \mathcal{O}(n)$  and  $r_n = \mathcal{O}(n)$ . Therefore,  $G_T(u) - G_T(0) = \Phi(u) - \Phi(0) + \mathcal{O}(\frac{1}{\sqrt{n}})$ , where  $\Phi$  is the standard Gaussian CDF. The exponential factor in (32) suggests that for sufficiently large  $n$  we only need a good approximation of  $\Phi(u) - \Phi(0)$  around 0, and only within a small fraction of the standard deviation of  $T_n$ ,  $\sqrt{\ddot{\mu}_T^n(s)}$ . Thus, it is sufficient to use the following bounds:  $1 - \frac{u^2}{2} \leq e^{-\frac{u^2}{2}} \leq 1$  which produces, if  $u \geq 0$ ,  $\frac{u}{\sqrt{2\pi}} - \frac{u^3}{6\sqrt{2\pi}} \leq \Phi(u) - \Phi(0) \leq \frac{u}{\sqrt{2\pi}}$ . Then, it is easy to show that

$$\int_{\dot{\mu}_T^n(s)}^{\infty} Q_T(t) e^{-s(t - \dot{\mu}_T^n(s))} dt = \frac{1}{\sqrt{2\pi s^2 \ddot{\mu}_T^n(s)}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

Replacing this in (31) produces (23), the first result of the theorem. The computation of  $\mathbb{P}(T_n < \tau_n)$  is similar and it is omitted.

### B. Proof of Corollary 1

In order to apply Thm. 1, the Gaussian quadratic form (25) can be expressed as the sum of  $n'$  independent variables. Then,

we only need to evaluate its LMGF  $\mu_i(s) = \log \mathbb{E}_i(e^{sT_{n'}})$  under  $\mathcal{H}_i$ ,  $i = 0, 1$ , and their derivatives:

$$\mu_i(s) = -\log \det(I_n - s\Xi_{i,n'}P_{n'}), \quad (33)$$

$$\dot{\mu}_i(s) = \text{tr}\{[(\Xi_{i,n'}P_{n'})^{-1} - sI_{n'}]^{-1}\}, \quad (34)$$

$$\ddot{\mu}_i(s) = \text{tr}\{[(\Xi_{i,n'}P_{n'})^{-1} - sI_{n'}]^{-2}\}. \quad (35)$$

Replacing these expressions in (23) and (24) completes the proof.

## APPENDIX II

### PERFORMANCE FOR THE EQUICORRELATED PROCESS: PROOF OF LEMMA 1

#### A. $P_{fa}$ for the LCS strategy

We use Th. 1 to compute  $P_{fa,LCS}$ . We first evaluate the LMGF of the statistic  $T_{LCS}$  in (20) under  $\mathcal{H}_0$  and its first and second derivatives given by (33)-(35) in Corollary 1, where  $P_{n'} = \Gamma_{n'}$ .

The covariance matrix of the equicorrelated process is a circulant matrix and, therefore, it is diagonalized by the  $l \times l$  DFT matrix  $F_{ll}$  and its eigenvalues are the DFT of its first column, i.e.,  $\Sigma_{s,l} = F_{ll}D_{s,l}F_{ll}^H$  where  $D_{s,l} = \text{diag}(1 + (l-1)\rho, 1-\rho, \dots, 1-\rho)$ . Considering the PSD of the equicorrelated process, the frequency vectors used in the precoding matrix of the PFS-MAC scheme given in Def. 2 have the indexes  $j_{k'} = k' - 1$ ,  $k' = 1, \dots, l'$ . Then, the covariance matrix in (13) is

$$\Xi_{1,l'} = \gamma_0^2 D_{s,l'} + (\gamma_0^2 \sigma_v^2 + \sigma_w^2) I_{l'},$$

where  $D_{s,l'}$  is a  $l' \times l'$  diagonal matrix,  $D_{s,l'} = \text{diag}(1 + (l-1)\rho, 1-\rho, \dots, 1-\rho)$ . Notice that  $\Xi_{1,l'}$  is a diagonal matrix because the covariance of the process is a circulant matrix and the precoding scheme is based on the DFT matrix. The covariance matrices of  $z$  under  $\mathcal{H}_0$  and  $\mathcal{H}_1$  are given by (15) and they result

$$\begin{aligned} \Xi_{0,n'} &= (\gamma_0^2 \sigma_v^2 + \sigma_w^2) I_{n'}, \\ \Xi_{1,n'} &= \gamma_0^2 \sigma_s^2 (1-\rho) I_{n'} + \gamma_0^2 \sigma_s^2 \rho (I_m \otimes F_{ll'}^H) \mathbf{1}_l (I_m \otimes F_{ll'}) \\ &\quad + (\gamma_0^2 \sigma_v^2 + \sigma_w^2) I_{n'}, \\ &= \gamma_0^2 \sigma_s^2 ((1-\rho) I_{n'} + \rho \mathbf{1}_m \otimes E_{l'}) + (\gamma_0^2 \sigma_v^2 + \sigma_w^2) I_{n'}, \end{aligned}$$

where  $\otimes$  is the Kronecker product and the  $l' \times l'$  matrix  $E_{l'}$  is defined as  $E_{l'} = \text{diag}(l, 0, \dots, 0)$ . Then,

$$\Xi_{0,n'} \Gamma_{n'} = I_m \otimes \text{diag}\left(\frac{\gamma_0^2 d_1}{\gamma_0^2 (d_1 + \sigma_v^2) + \sigma_w^2}, \dots, \frac{\gamma_0^2 d_{l'}}{\gamma_0^2 (d_{l'} + \sigma_v^2) + \sigma_w^2}\right),$$

where  $d_1 = \sigma_s^2(1 + (l-1)\rho)$  and  $d_k = \sigma_s^2(1-\rho)$  if  $k = 2, \dots, l'$ . Using (33)-(35) of Corollary 1, and denoting  $\Gamma = \text{SNR}_{\text{PFS-MAC}}(l)$ , the LMGF of the statistic (20) under  $\mathcal{H}_0$  and

its first and second derivatives are:

$$\begin{aligned} \mu_0(s_0) &= -\frac{n}{l} \log\left(1 - s_0 \frac{\Gamma(1+(l-1)\rho)}{1+\Gamma(1+(l-1)\rho)}\right) \\ &\quad - n\left(\beta - \frac{1}{l}\right) \log\left(1 - s_0 \frac{\Gamma(1-\rho)}{1+\Gamma(1-\rho)}\right), \end{aligned} \quad (36)$$

$$\begin{aligned} \dot{\mu}_0(s_0) &= \frac{n}{l} \frac{\Gamma(1+(l-1)\rho)}{1+\Gamma(1+(l-1)\rho) - s_0 \Gamma(1+(l-1)\rho)} \\ &\quad + n\left(\beta - \frac{1}{l}\right) \frac{\Gamma(1-\rho)}{1+\Gamma(1-\rho) - s_0 \Gamma(1-\rho)}, \end{aligned} \quad (37)$$

$$\begin{aligned} \ddot{\mu}_0(s_0) &= \frac{n}{l} \frac{(\Gamma(1+(l-1)\rho))^2}{(1+\Gamma(1+(l-1)\rho) - s_0 \Gamma(1+(l-1)\rho))^2} \\ &\quad + n\left(\beta - \frac{1}{l}\right) \frac{(\Gamma(1-\rho))^2}{(1+\Gamma(1-\rho) - s_0 \Gamma(1-\rho))^2}. \end{aligned} \quad (38)$$

Now, the mean of the statistic  $T_{LCS}$  under  $\mathcal{H}_1$  is  $m_1 = \mathbb{E}_1(T_{LCS}) = n\Gamma(\rho + \beta(1-\rho))$ . Then,  $\mu_0(s_0^*) = \tau_n = m_1$  implies that  $s_0^* = 1$ . Replacing it in (36)-(38) and using Th. 1, we obtain (27).

#### B. $P_{fa}$ for the FCS strategy

The statistic  $T_{FCS}$  is an LLR, for which the LMGFs under  $\mathcal{H}_0$  and  $\mathcal{H}_1$  satisfy

$$\mu_0(s) = \mu_1(s-1), \quad (39)$$

where  $\mu_i(s) = \mathbb{E}_i(e^{sT_{FCS}})$ . The direct computation of  $\mu_0(s)$  implies to obtain firstly the inverse of  $\Xi_{1,n'}$ , which is not straightforward. An easier way is to compute  $\mu_1(s)$  for which we only need to invert the diagonal matrix  $\Xi_{0,n'}$  and then, to use (39) to evaluate  $\mu_0(s)$ . The expression for  $\mu_1(s)$  and its first and second derivatives are:

$$\begin{aligned} \mu_1(s_1) &= -s_1 \log(1 + \Gamma(1 + (n-1)\rho)) - s_1(n\beta - 1) \log(1 + \Gamma(1-\rho)) \\ &\quad - \log(1 - s_1 \Gamma(1 + (n-1)\rho)) - (n\beta - 1) \log(1 - s_1 \Gamma(1-\rho)), \\ \dot{\mu}_1(s_1) &= -\log(1 + \Gamma(1 + (n-1)\rho)) - (n\beta - 1) \log(1 + \Gamma(1-\rho)) \\ &\quad + \frac{\Gamma(1 + (n-1)\rho)}{1 - s_1 \Gamma(1 + (n-1)\rho)} + (n\beta - 1) \frac{\Gamma(1-\rho)}{1 - s_1 \Gamma(1-\rho)}, \\ \ddot{\mu}_1(s_1) &= \frac{(\Gamma(1 + (n-1)\rho))^2}{(1 - s_1 \Gamma(1 + (n-1)\rho))^2} + (n\beta - 1) \frac{(\Gamma(1-\rho))^2}{(1 - s_1 \Gamma(1-\rho))^2}. \end{aligned}$$

$\tau_{FCS} = \mathbb{E}_1(T_{FCS}) = \dot{\mu}_1(s_1) = \dot{\mu}_0(s_0)$  implies that  $s_1 = 0$  and  $s_0 = 1$ . Using this in (39) and replacing all in (23), we obtain (28). Notice that this result coincides with the one in (27) if  $l$  is replaced by  $n$ . The difference between both FCS and LCS strategies is that the first one contemplates the full correlation of the process.

#### C. $P_{fa}$ for the ED strategy

The energy detector  $T_{ED}$  can be normalized by  $\gamma_0^2 \sigma_v^2 + \sigma_w^2$  without modifying the performance:  $\bar{T}_{ED} = \frac{T_{ED}}{\gamma_0^2 \sigma_v^2 + \sigma_w^2}$ . Its LMGF under  $\mathcal{H}_0$  and first and second derivatives are:

$$\mu_0(s_0) = -n \log(1 - s_0), \quad (40)$$

$$\dot{\mu}_0(s_0) = \frac{n}{1 - s_0}, \quad (41)$$

$$\ddot{\mu}_0(s_0) = \frac{n}{(1 - s_0)^2}. \quad (42)$$

$\tau_{ED} = \mathbb{E}_1(\tilde{T}_{ED}) = n(1 + \Gamma) = \dot{\mu}_0(s_0)$  implies that  $s_0 = \frac{\Gamma}{1+\Gamma}$ . Using this in (40)-(42) and replacing all in (23) results in (29). Notice that the same result can be obtained by replacing  $l$  by  $n$  in (27) and taking  $\rho = 0$ , given that the ED strategy does not consider the correlation of the process.

## REFERENCES

- [1] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 100–117, 2012.
- [2] R. Tenney and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, no. 4, pp. 501–510, July 1981.
- [3] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I: Fundamentals," in *Proc. IEEE*, vol. 85, Jan. 1997, pp. 54–63.
- [4] J. N. Tsitsiklis, "Decentralized detection," *Adv. Stat. Signal Process.*, vol. 2, pp. 297–344, Sep. 1993.
- [5] J.-F. Chamberland and V. Veeravalli, "Decentralized detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.
- [6] A. D'Costa, V. Ramachandran, and A. Sayeed, "Distributed classification of Gaussian space-time sources in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1026–1036, Aug 2004.
- [7] J.-F. Chamberland and V. Veeravalli, "How dense should a sensor network be for detection with correlated observations?" *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5099–5106, Nov 2006.
- [8] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1007–1015, 2004.
- [9] W. Li and H. Dai, "Distributed Detection in Wireless Sensor Networks Using A Multiple Access Channel," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 822–833, 2007.
- [10] K. Cohen and A. Leshem, "Performance analysis of likelihood-based multiple access for detection over fading channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2471–2481, 2013.
- [11] P. Bianchi, J. Jakubowicz, and F. Roueff, "Linear Precoders for the Detection of a Gaussian Process in Wireless Sensors Networks," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 882–894, March 2011.
- [12] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: distributed detection of a known signal in dependent Gaussian noise," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3266–3279, 2000.
- [13] H. Chen, B. Chen, and P. Varshney, "A new framework for distributed detection with conditionally dependent observations," *Signal Processing, IEEE Transactions on*, vol. 60, no. 3, pp. 1409–1419, March 2012.
- [14] J. Maya, L. Rey Vega, and C. Galarza, "Optimal Resource Allocation for Detection of a Gaussian Process Using a MAC in WSNs," *Signal Processing, IEEE Transactions on*, vol. 63, no. 8, pp. 2057–2069, April 2015.
- [15] Y. Sung, H. Poor, and H. Yu, "How much information can one get from a wireless ad hoc sensor network over a correlated random field?" *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2827–2847, June 2009.
- [16] J. Unnikrishnan and V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 1, pp. 18–27, Feb 2008.
- [17] Y. Sung, S. Misra, L. Tong, and A. Ephremides, "Cooperative routing for distributed detection in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 471–483, February 2007.
- [18] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov Random Fields With Nearest-Neighbor Dependency," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 816–827, Feb 2009.
- [19] H. He, A. Subramanian, P. Varshney, and T. Damarla, "Fusing heterogeneous data for detection under non-stationary dependence," in *Information Fusion (FUSION), 2012 15th International Conference on*, July 2012, pp. 1792–1799.
- [20] V. R. Aalo, V., "On distributed detection with correlated sensors: two examples," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 3, pp. 414–421, May 1989.
- [21] E. Drakopoulos and C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593–606, Jul 1991.
- [22] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys & Tutorials, IEEE*, vol. 11, no. 1, pp. 116–130, 2009.
- [23] J. Lunden, V. Koivunen, and H. Poor, "Spectrum exploration and exploitation for cognitive radio: Recent advances," *Signal Processing Magazine, IEEE*, vol. 32, no. 3, pp. 123–140, May 2015.
- [24] T.-Y. Wang, Y. S. Han, P. K. Varshney, and P.-N. Chen, "Distributed fault-tolerant cation in wireless sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 4, pp. 724–734, 2005.
- [25] S. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*, 1st ed. Prentice-Hall, 1998.
- [26] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*, 2nd ed. Wiley, 2009.
- [27] J. Maya, C. Galarza, and L. Rey Vega, "Exploiting Spatial Correlation in Energy Constrained Distributed Detection," *ArXiv*, Sept. 2015.
- [28] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters*, vol. 27, no. 23, pp. 2145–2146, Nov 1991.
- [29] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.
- [30] W. Feller, *An Introduction to Probability Theory and its Applications*. New York, NY, USA: John Wiley & Sons, Inc., 1966, vol. 2.